CONDITIONING OF TWO-RESPONSE PATTERNS OF KEY PECKING IN PIGEONS[1]

REBECCA J. GRAYSON AND EDWARD A. WASSERMAN

THE UNIVERSITY OF IOWA

On discrete trials, two response keys were made available to hungry pigeons and food reinforcement depended on the order in which the required two key pecks occurred. In different phases, only one of the four possible two-peck sequences (left-left, left-right, right-left, and right-right) produced food reinforcement. In each case, the pigeons learned to perform the correct two-peck sequence more often than the incorrect sequences. Furthermore, the course of differentiation mastery indicated that both reinforcement history and response-reinforcer contiguity influenced performance. These results reveal that response patterns comprising two instances of the same response left-left and right-right) or instances of two different responses (left-right and right-left) may function as operants, thereby extending the generality of conditioning principles from discrete responses to structured sequences of behavior. These and other results are discussed in terms of contiguity-based and memory-based models of learning.

Key words: response differentiation, response patterning, procedural-functional operant, response-reinforcer contiguity, short-term memory, key peck, pigeons

Procedurally, an operant is a class of behaviors on which reinforcement depends (Schick, 1971). Functionally, an operant is a class of behaviors whose frequency of occurrence is increased when responses of that class are required to produce reinforcement (Zeiler, 1977).

Experimenters have traditionally employed discrete responses such as key pecks and lever presses as procedural operants. However, researchers have recently begun to explore the possibility of conditioning operants that entail instances of two or more discrete responses. Here, the selective reinforcement of inter-response times (Platt, 1973) is to be noted, as is the selective reinforcement of increasing and decreasing rates of responding within brief spans of time (Hawkes and Shimp, 1975; Wasserman, 1977). The results of these investigations encourage the view that the principles of operant conditioning have applicability to organized patterns of behavior (Shimp, 1976a).

In the above examples of multiresponse operants, the patterns of behavior involved repeated instances of the *same* response at different points in time. It should also be possible to condition behavioral patterns comprising instances of two or more *different* responses. This is the focus of the present paper.

Historically, research and theory in the acquisition of response sequences have centered on the concept of chaining (Skinner, 1938). Here, the essential notion is that elaborate chains of behavior can be experimentally constructed by adding topographically different responses to the beginning of already established response sequences (but see Weiss, 1978). Thus, after first requiring an organism to perform Response 1 in order to obtain reinforcement, it is now required to perform the sequence Response 2-Response 1. After mastery of the Response 2-Response 1 sequence, training on the Response 3-Response 2-Response 1 sequence ensues, and so on until the criterion sequence has been conditioned.

Although relevant to the issue of response-sequence learning, studies of response chaining are much less relevant to the question of whether organized sequences of responses can be differentiable operants. Minimally, in order to show that response differentiation has successfully occurred requires (a) that reinforce-

ment be made dependent on a subset of all possible response variants and (b) that a selective increase in the frequency of that subset above all other variants occurs as a consequence of differential reinforcement (Notterman and Mintz, 1965). Thus, investigations of response chaining as typically conducted (e.g., Sidman and Rosenberger, 1967) cannot specify whether response-pattern differentiation has taken place because incorrect response sequences are not permitted to proceed to completion. In the Response 2-Response 1 phase of the preceding example, any sequence beginning with Response 1 would ordinarily terminate the trial and produce a brief timeout or other aversive consequence. It then becomes impossible to monitor the frequency of occurrence of all possible response variants so that response-pattern differentiation might unequivocally be demonstrated.

The inability to monitor the incidence of all possible response variants also plagues methodologies that should support response-sequence learning, but that do not interrupt incorrect response sequences. An example of this method is found in the study of Catania (1971; tandem schedules). Here, following the passage of a variable time interval, pigeons were required to peck the left key and then the right key to receive reinforcement; pecks preceding the variable-interval requirement were of no consequence, nor were those that occurred before the left-right sequence after the variable-interval requirement. Thus, after the variable time interval had elapsed, the pigeon could respond right-right-left-left-right before reinforcer delivery (as one of a great many possibilities), or it could more efficiently respond left-right. In either case, reinforcement followed the left-right sequence, but how and when are we to determine the incidence of all other two-response sequences so as to verify the occurrence of response-pattern differentiation?

To remedy this problem of prior work, the present experiment required food-deprived pigeons to make two key pecks to receive grain reinforcement. Two response keys were made available and reinforcement after two responses depended on the order in which the keys were pecked. For instance, in one phase of training, food reinforcement followed the pattern of pecks to the two keys, left-right; the patterns right-left, left-left, and right-right failed to produce reinforcement, but were recorded.

If such selective reinforcement of a two-response pattern led to an increase in the frequency of occurrence of that particular pattern above all others, then this finding would support the conclusion that a succession of two different responses could be differentiated from the remaining sequences and that this two-response pattern could function as an operant. Successful differentiation of two-response patterns would further indicate that response-sequence learning need not proceed in a piecemeal fashion, but might instead begin with the criterion sequence, so long as its base level of occurrence were sufficiently high to produce occasional reinforcement.

METHOD

Subjects

Two experimentally naive homing pigeons were housed individually, with water and grit constantly available in their home cages. They were kept at 80% of their free-feeding weights by postsession feeding in the home cage.

Apparatus

The side keys of two identical (31.4 by 38.5 by 42.0 cm) Scientific Prototype (Model B200) three-key pigeon chambers were used. These 3.1-cm diameter keys had a minimum force requirement of 0.20 N for operation and could be transilluminated with red light from small bulbs (CM 1820, 24 V dc). The solenoid-actuated feeder was located directly below the center key and was illuminated with white light from a small bulb (ESB 24 V dc) when food was available. Chamber illumination was provided by a houselight (25 W, 110 V ac) mounted outside the chamber. To mask extraneous noises, each pigeon chamber was located in a separate room in which white noise was present throughout each session. A Digital Equipment Corporation PDP 8/L minicomputer, equipped with the SKED software system (Snapper, Stephens, and Lee, 1974), was located in another room and was used to control all procedures and to record all data for subsequent analysis.

Procedure

Preliminary training. The pigeons were first trained to eat grain from the lighted magazine when it was activated. The birds were then autoshaped to peck the left and right keys when they were singly illuminated in a ran-

dom order (Brown and Jenkins, 1968). A trial began with the illumination of one key. After 8 sec, the keylight was extinguished and the food hopper was raised for 2.5 sec. Intertrial intervals (ITIs) were equiprobably set at 30, 60, or 90 sec. There were 25 trials in each of three daily sessions, and by the third day, both pigeons were pecking the lighted keys on most of the trials.

In the next phase of preliminary training, the birds were required to make two key pecks to receive food. Here, a trial began with illumination of one of the two keys, equiprobably determined. The light was extinguished when that key was pecked. After 1 sec, one of the keys was again illuminated equiprobably. A peck on this key extinguished the light and presented grain reinforcement. Thus, each of the 50 trials within a daily session consisted of one of four equally probable sequences: left-left (LL), left-right (LR), right-left (RL), and right-right (RR). Trials were separated by 15-sec ITIs, during which both response keys were dark. This procedure lasted four days.

In the next session, the procedure was altered as follows: a trial began with illumination of both keylights. After two pecks, the keylights were extinguished. If the sequence of two pecks was LR or RL, 2.5 sec of food reinforcement followed; if the sequence was LL or RR, no reinforcement followed, but 2.5 sec were added to the 15-sec ITI. A session ended after 25 reinforcers or after 1 hr, whichever occurred first. This stage of training terminated when all 25 reinforcers were collected within 1 hr. For Bird 369, this required three sessions, for Bird 370 six sessions. This procedure preceded experimental training in order to reduce the anticipated tendency of the pigeons to peck twice on the same key, rather than switching from one key to the other. On the last day of pretraining, Birds 369 and 370 emitted the LL, LR, RL, and RR sequences 12, 9, 16, and 11 times and 27, 19, 6, and 70 times, respectively.

Experimental training. Here, the procedure was identical to the final portion of preliminary training, except that each bird was required to emit only one particular response sequence to produce reinforcement. For example, during a particular phase, the sequence LR was the only one followed by 2.5 sec reinforcement; all others caused 2.5 sec to be added to the 15-sec ITI. Sessions ended after 25 re-

inforcements or 1 hr, whichever came first. A bird reached criterion on a particular sequence when each of the other sequences occurred 12 or fewer times in a daily session. When criterion was reached, the required sequence was changed in the next session. The sequence of five phases and the sessions to reach criterion (in parentheses) were for Bird 369, LR(15), RL(12), RR(2), LL(2), and LR(20); for Bird 370, LR(2), RL(23), LL(2), RR(3), and RL(22).

The number of times that each response sequence was performed in each session of experimental training served as the principal measure of performance.

RESULTS

Both birds learned to perform the response sequence that produced reinforcement more than all of the other incorrect sequences. Figure 1 shows the performance of each bird on each criterion day in the five experimental phases. The shaded bar in each phase represents the response sequence that produced reinforcement. In comparing the frequencies of the other sequences, two trends can be discerned. First, in those phases that required LR and RL sequences, the next most frequently emitted sequence was RR and LL, respectively; these incorrect sequences involved two pecks to the key most contiguous with reinforcement. Second, in those phases that required LL and RR sequences, the next most frequently emitted sequence was usually the one that had led to reinforcement in the prior phase. The exception was Bird 369 in Phase III.

The daily performance of each bird is plotted in Figure 2. Each point on the individual functions corresponds to the number of times that a particicular sequence occurred in a session. During the phases that required either LR or RL for reinforcement, the sequence involving two pecks on the key most contiguous with reinforcement (RR and LL, respectively) was the last to reach criterion level. In the first sessions of a phase that required either LR or RL, the performance of the birds was noticeably influenced both by the prevailing reinforcement dependency and by that of the prior phase. For example, in Session 16, Bird 369 was required to perform RL after having reached criterion on LR the day before. Occurrences of LR were quite high in Session 16 and, even though the pigeon had received rein-
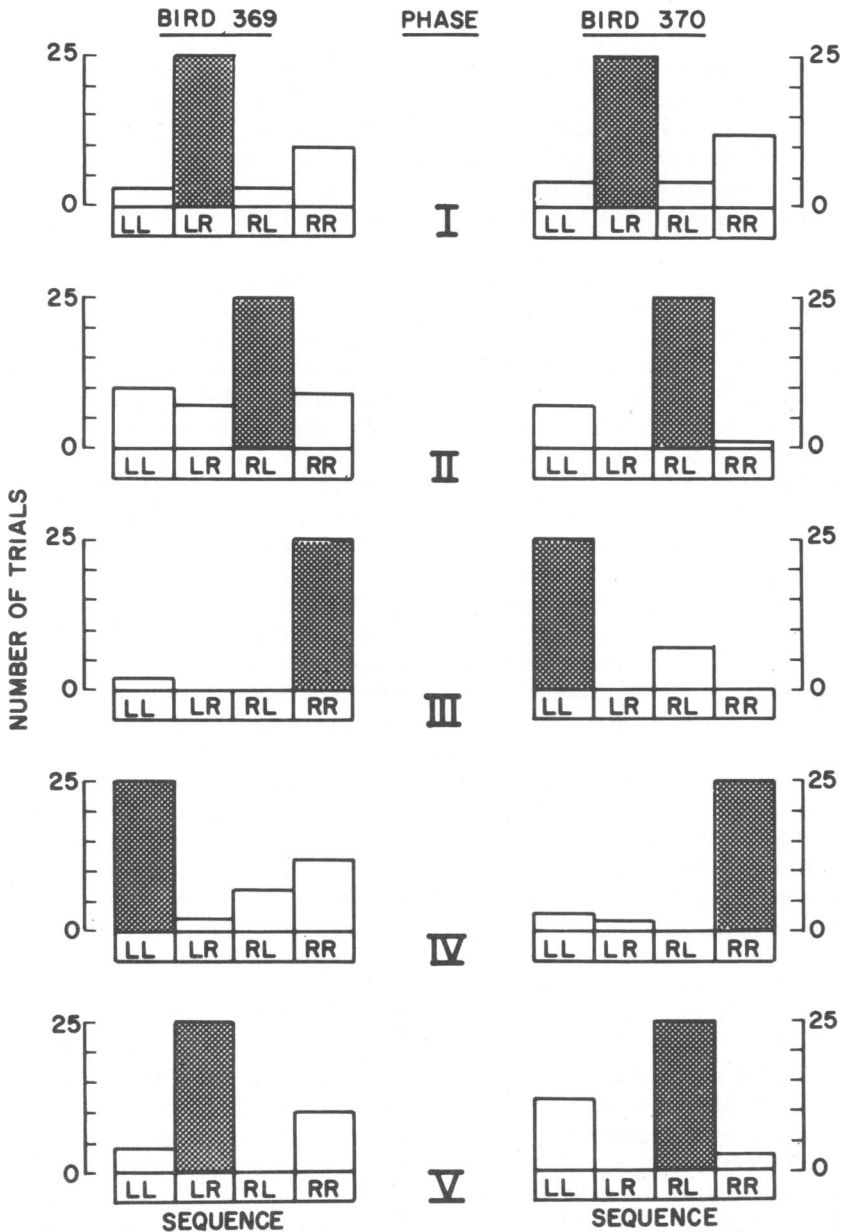
Fig. 1. Number of trials on which the LL, LR, RL, and RR sequences occurred on the criterion day of Phases I, II, III, IV, and V for Birds 369 and 370. The shaded bars represent the two-response sequences that produced reinforcement.

forcement after the RL sequence only five times in the session, there were many occurrences of the LL sequence. Finally, Figure 2 shows that learning of the LL and RR sequences generally occurred much more rapidly than learning of the LR and RL sequences. Here, the exception was Bird 370 learning the LR sequence in Phase I.

DISCUSSION

These data show that the sequence of pecking two keys is a differentiable property of the pigeon's behavior. In each experimental phase, the frequency of the correct response sequence increased and that of all other sequences decreased to criterion levels.
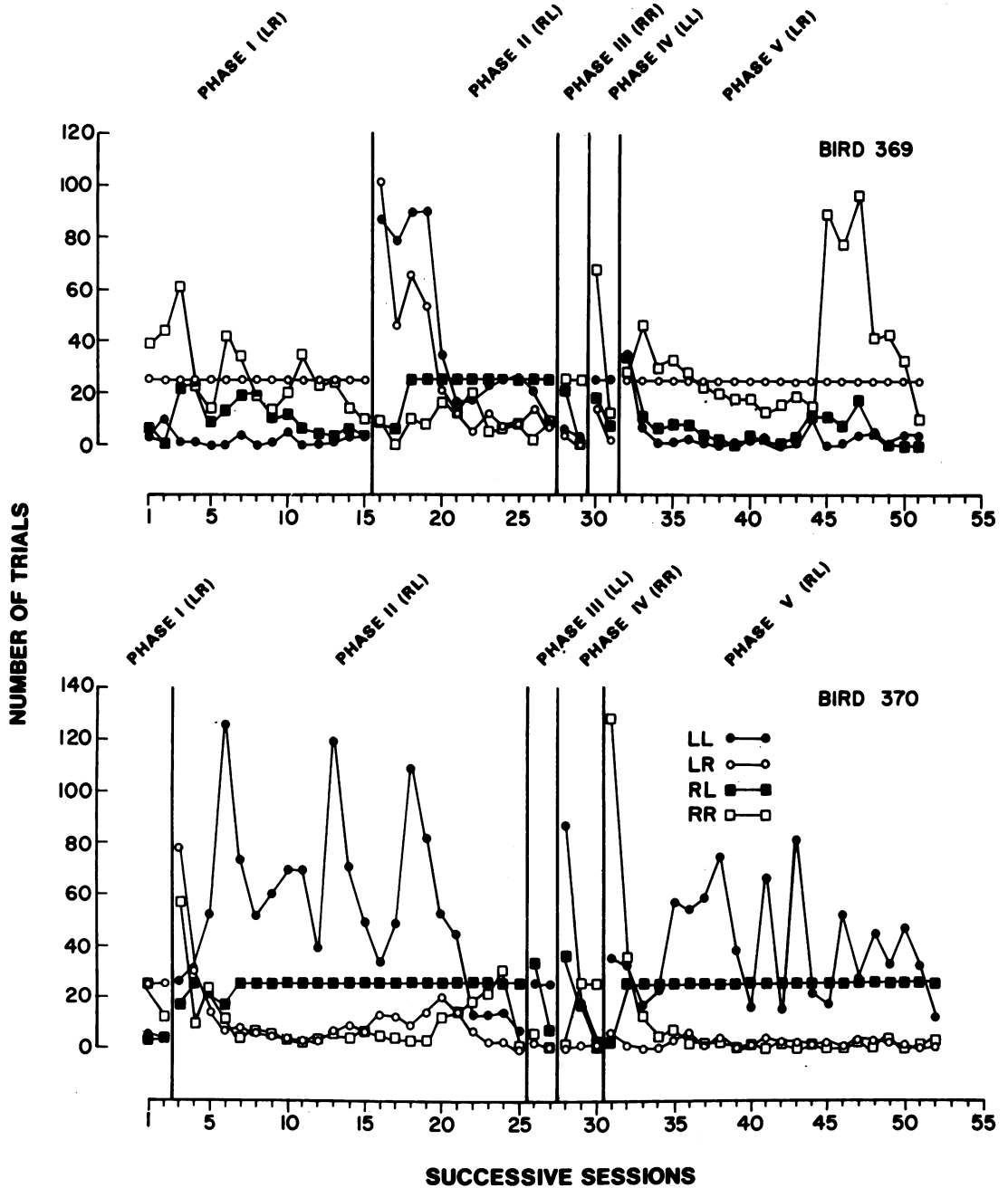
Fig. 2. Number of trials on which the LL, LR, RL, and RR sequences occurred on each day of experimental training for Birds 369 and 370. Vertical lines separate the five experimental phases that are labelled according to the two-response sequence that produced reinforcement.

Of special interest was criterion performance in those phases that required pecks to two different response keys (LR and RL sequences). Here, the most frequent error involved two pecks to the key most contiguous with food reinforcement. This result is consistent with

accounts of responding that stress response-reinforcer contiguity.

Catania (1971) applied the notion of response-reinforcer contiguity to schedules requiring a sequence of two different responses for reinforcement. Specifically, Catania studied

the effects of reinforcement following various two-key response sequences on the pigeon's relative rates of responding to the two keys. He found that when particular response sequences (*e.g.*, LR, LRR, LRRR) produced reinforcement, the percentage of pecks to each key was related to the distance of left and right pecks from the reinforcer. Catania concluded that each peck preceding reinforcement makes an independent contribution to later responding, and that the size of that contribution is an inverse function of the delay of reinforcement.

Consistent with Catania's predictions, more responses occurred in the present study to the key most proximal to reinforcer delivery than to the other key in those phases that required pecks to two different keys. The overall percentage of pecks to the last key in the two-peck sequence on the criterion days of Phases I, II, and V for Birds 369 and 370 was, respectively, 58.54, 50.98, 57.69, and 58.89, 59.09, 61.25 (see Figure 1 for the raw data).

Despite the plausibility of Catania's explanation of relative response output, his contiguity-based account is of little value in explaining the differentiations of the four possible two-peck sequences that we obtained. Suppose that the sequence producing reinforcement is LR. If a peck to either key in either position of a two-peck sequence is determined solely by response-reinforcer contiguity, then the probability of the two-peck sequences should be ordered as follows: $p(LL) < p(LR) = p(RL) < p(RR)$.[2] This ordering is clearly at odds with the present data, which generally yielded $p(RL) = p(LL) < p(RR) < p(LR)$. So too are variations on the theme of response-reinforcer contiguity that take into account the tendency of pigeons to continue to peck the key where they begin. Depending on the strength of the bias to continue pecking the same key, the following orderings are expected: $p(RL) < p(LL) < p(LR) < p(RR)$ and $p(RL) < p(LR) < p(LL) < p(RR)$.[3]

The above versions of contiguity theory assume that both responses of the two-peck sequences are equally memorable at the time of reinforcement, and that the effect of response-

reinforcer delay is simply to increase the strength of responding to each key to varying degrees. Accounts stressing short-term memory question the first assumption of the contiguity formulation, and instead propose that the effects of response-reinforcer delay are mediated by memory loss.

One prominent memory-based account of operant conditioning is that of Shimp (1976a). He suggested that a functional behavioral unit involves the contents of the organism's short-term memory at the moment of reinforcement. Shimp's research (Shimp, 1976b; Shimp and Moffitt, 1974) has revealed that the pigeon can remember as many as three responses, thus supporting the view that a functional unit may comprise more than a single response. Shimp's research has further found that the organism's memory for its own behavior fades the longer it has been since a response occurred. Extending these facts to the present data, we can appreciate how response differentiation is possible for behavioral units comprising two responses: an ordered sequence of two responses should still be available in the pigeon's short-term memory at the time of reinforcement or nonreinforcement. And we can understand why incorrect sequences often involve the response closest in time to reinforcement: memory for the second response of the positive sequence should exceed that of the first, thereby favoring association of the final response element of the sequence with reinforcement over association of the entire two-response sequence with reinforcement.

Of course, some would take issue with the introduction of the memory construct into the experimental analysis of behavior (*e.g.*, Branch, 1977). The present data might thus be alternatively interpreted as reflecting the process of response generalization. For example, if the RR sequence were more similar to the LR sequence than were RL and LL sequences, then the error patterns observed in Phases I, II, and V might not encourage the postulation of mem-

---

[2] Assume that following LR with reinforcement leads to $p(L) = 0.40$ and $p(R) = 0.60$. Then, $p(LL) = (0.40) \times (0.40) = 0.16$, $p(LR) = (0.40) \times (0.60) = 0.24$, $p(RL) = (0.60) \times (0.40) = 0.24$, and $p(RR) = (0.60) \times (0.60) = 0.36$.

[3] The first ordering is produced when a factor of 1.20 is assigned to the probability of pecking the same key as was first pecked. Thus, $p(LL) = (0.40) \times (0.48) = 0.19$, $p(LR) = (0.40) \times (0.52) = 0.21$, $p(RL) = (0.60) \times (0.28) = 0.17$, and $p(RR) = (0.60) \times (0.72) = 0.43$. The second ordering is produced when a factor of 1.40 is assigned to the probability of pecking the same key as was first pecked. Thus, $p(LL) = (0.40) \times (0.56) = 0.22$, $p(LR) = (0.40) \times (0.44) = 0.18$, $p(RL) = (0.60) \times (0.16) = 0.10$, and $p(RR) = (0.60) \times (0.84) = 0.50$.

ory processes. While this possibility is not to-
tally excluded by the present data, the results
of Phases III and IV reduce the plausibility
of this alternative account. If the LR sequence
is most similar to the RR sequence, and if the
RL sequence is most similar to the LL se-
quence (as would be required to explain the six
out of six positive instances of Phases I, II, and
V), then in Phases III and IV, sequences LR
and RL, respectively, should have been the
most frequent errors when the correct se-
quences were RR and LL. However, Figure 1
shows that this outcome occurred only one time
out of four.

Whether one's theoretical predilections favor
or oppose the construct of memory, it seems
clear that increasing importance must be as-
signed to processes capable of bridging the time
intervals that typically intervene between the
elements of learned associations (Medin, Rob-
erts, and Davis, 1976). Although research in the
area of short-term memory has historically
stressed the retention of externally-presented
stimuli, there is no reason to suppose that simi-
lar memory processes may not also be involved
in the retention of subject-generated responses
(*e.g.,* Maki, Moe, and Bierley, 1977; Shimp,
1976*b;* Shimp and Moffitt, 1974; Ziriax and
Silberberg, 1978). The greater success of a
memory-based model of operant behavior over
a more traditional contiguity-based formula-
·tion in explaining our own experimental find-
ings should encourage others to consider the
usefulness of the memory construct to the ex-
perimental analysis of behavior (for a recent
example of the usefulness of memory in un-
derstanding schedule control see Stubbs, Vau-
tin, Reid, and Delehanty, 1978).

REFERENCES

Branch, M. N. On the role of "memory" in the analy-
sis of behavior. *Journal of the Experimental Analysis
of Behavior,* 1977, **28,** 171-179.
Brown, P. L. and Jenkins, H. M. Auto-shaping of the
pigeon's key-peck. *Journal of the Experimental
Analysis of Behavior,* 1968, **11,** 1-8.
Catania, A. C. Reinforcement schedules: the role of
responses preceding the one that produces the rein-
forcer. *Journal of the Experimental Analysis of Be-
havior,* 1971, **15,** 271-287.

Hawkes, L. and Shimp, C. P. Reinforcement of be-
havioral patterns: shaping a scallop. *Journal of the
Experimental Analysis of Behavior,* 1975, **23,** 3-16.
Maki, W. S., Jr., Moe, J. C., and Bierley, C. M. Short-
term memory for stimuli, responses, and reinforcers.
*Journal of Experimental Psychology: Animal Be-
havior Processes,* 1977, **3,** 156-177.
Medin, D. L., Roberts, W. A., and Davis, R. T. (Eds)
Processes of animal memory. Hillsdale, New Jersey:
Erlbaum, 1976.
Notterman, J. M. and Mintz, D. E. *Dynamics of re-
sponse.* New York: Wiley, 1965.
Platt, J. R. Percentile reinforcement: paradigms for
experimental analysis of response shaping. In G.
H. Bower (Ed), *The psychology of learning and
motivation: advances in research and theory,* Vol. 7.
New York: Academic Press, 1973. Pp. 271-296.
Schick, K. Operants. *Journal of the Experimental
Analysis of Behavior,* 1971, **15,** 413-423.
Shimp, C. P. Organization in memory and behavior.
Journal of the Experimental Analysis of Behavior,
1976, **26,** 113-130. (*a*)
Shimp, C. P. Short-term memory in the pigeon: rela-
tive recency. *Journal of the Experimental Analysis
of Behavior,* 1976, **25,** 55-61. (*b*)
Shimp, C. P. and Moffitt, M. Short-term memory in
the pigeon: stimulus-response associations. *Journal
of the Experimental Analysis of Behavior,* 1974, **22,**
507-512.
Sidman, M. and Rosenberger, P. B. Several methods
for teaching serial position sequences to monkeys.
Journal of the Experimental Analysis of Behavior,
1967, **10,** 467-478.
Skinner, B. F. *The behavior of organisms.* New York:
Appleton-Century-Crofts, 1938.
Snapper, A. G., Stephens, K. R., and Lee, D. M. *The
SKED software system.* Kalamazoo, Michigan: The
SKED Users Group, 1974.
Stubbs, D. A., Vautin, S. J., Reid, H. M., and Delehanty,
D. L. Discriminative functions of schedule stimuli
and memory: a combination of schedule and choice
procedures. *Journal of the Experimental Analysis of
Behavior,* 1978, **29,** 167-180.
Wasserman, E. A. Conditioning of within-trial pat-
terns of key pecking in pigeons. *Journal of the Ex-
perimental Analysis of Behavior,* 1977, **28,** 213-220.
Weiss, K. M. A comparison of forward and backward
procedures for the acquisition of response chains in
humans. *Journal of the Experimental Analysis of
Behavior,* 1978, **29,** 255-259.
Zeiler, M. Schedules of reinforcement: the controlling
variables. In W. K. Honig and J. E. R. Staddon
(Eds), *Handbook of operant behavior.* Englewood
Cliffs, New Jersey: Prentice-Hall, 1977. Pp. 201-232.
Ziriax, J. M. and Silberberg, A. Discrimination and
emission of different key-peck durations in the pi-
geon. *Journal of Experimental Psychology: Animal
Behavior Processes,* 1978, **4,** 1-21.